

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: August 22, 2003, 13:48:21; Search time 31.5 Seconds

(without alignments)
4636.982 Million cell updates/sec

Title: US-09-745-506-74

Perfect score: 2789
Sequence: 1 GTGATTGTTACTTGTGCTCTGTTACTTAACATTCAA 1553

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=tlp
-Q=/cgn2_1/USPTO.spool/US09745506/runat_22082003_132709_10731/apz_query.fasta_1.1735
-DB=SwissProt_41 -QPMT=fastan -SUFFIX=n2p -MIMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=100 -THR_SCORE=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -THR_MAX=100 -MINLEN=0 -MAXLEN=2000000000
-USER=US09745506.ecgn_1.1.40.gnarat_22082003_132709_10731 -NCPU=6 -ICPU=3
-NO_MAP -LANG=ENGLISH -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1799	64.5	350	1	NF3L_HUMAN
2	1575	56.5	350	1	NF3L_MOUSE
3	578	20.7	292	1	A36A_DROME
4	453.5	16.3	278	1	YJ3C_SCHPO
5	384.5	13.8	288	1	NF3_YEAST
6	297	10.6	262	1	YK04_CLOST
7	285.5	10.2	373	1	YOF0_BACSU
8	278	10.0	366	1	YF59_STEAM
9	264	9.5	268	1	YD03_CLOAB
10	263.5	9.4	372	1	YD80_BACHD
11	256	9.2	373	1	YEB8_LISIN
12	240	8.6	373	1	YES2_LISMO
13	225	8.1	241	1	Y705_CAMJE
14	216.5	7.8	379	1	YM30_MYCTU
15	210	7.5	262	1	Y931_STRPY
16	209.5	7.5	263	1	YMI6_ANASP
17	208	7.5	285	1	YN01_STRCO
18	203.5	7.3	265	1	YG09_STRPC

19	199.5	7.2	385	1	YM30_MYCLE	069481 mycobacteri
20	198	7.1	262	1	Y351_UREPA	09960 ureaplasma
21	193	6.9	243	1	Y959_HELPU	02832 helicobacte
22	191	6.8	244	1	Y927_METUA	09837 methanococ
23	187.5	6.7	257	1	YA73_LACIA	025613 lactococcus
24	181	6.5	243	1	Y959_HELPU	025613 helicobacte
25	148	5.3	251	1	Y456_MYCPV	09890 mycoplasma
26	133	4.8	247	1	YBGI_SALTY	08347 salmonella
27	129	4.6	247	1	YBGI_ECOLI	075743 escherichia
28	117	4.2	248	1	Y468_BORBU	051424 borrelia bu
29	117	4.0	252	1	YK93_VIBCH	09349 vibrio chol
30	111	4.0	251	1	Y183_PASMU	09CP12 pasteurella
31	110	3.9	248	1	Y731_RALSO	08VA60 ralsosonia s
32	107	3.8	250	1	Y627_PYRHO	058361 pyrococcus
33	106	3.8	252	1	Y8E5_PSEAE	09HY42 pseudomonas
34	105.5	3.8	3829	1	SACS_HUMAN	09ZJ4 homo sapien
35	103.5	3.7	3830	1	SACS_MOUSE	09J1C8 mus musculu
36	103	3.7	674	1	XP22_HUMAN	043895 homo sapien
37	102.5	3.7	1153	1	A3D1_HUMAN	014617 homo sapien
38	102	3.7	249	1	Y382_NEIMA	09JW6 neisseria m
39	101	3.6	828	1	SC10_ARATH	081W3 arabidopsis
40	100	3.6	249	1	YK54_NEITB	09JX9 neisseria m
41	100	3.6	841	1	HEIX_METUA	057742 methanococ
42	98	3.5	1161	1	POL_SFVL	023074 simian foam
43	97.5	3.5	245	1	YH77_ARCFU	028497 archaeoglob
44	97.5	3.5	502	1	GSPE_PSEAE	000512 pseudomonas
45	97.5	3.5	3259	1	GTAN_HUMAN	014789 homo sapien

ALIGNMENTS

RESULT 1
NF3L_HUMAN
AC NF3L_HUMAN STANDARD; PRT; 350 AA.
ID Q9GZT8; Q9H2D2; Q9HC18;
DT 28-FEB-2003 (Ref. 41, Created)
DT 28-FEB-2003 (Ref. 41, Last sequence update)
DT 15-SEP-2003 (Ref. 42, Last annotation update)
DE NF3-like protein 1 (Amyotrophic lateral sclerosis 2 chromosomal
DE region candidate gene protein 1) (Miy018 protein) (MDS015).
GN NF3L1 OR ALS2CR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND SUBCELLULAR LOCATION.
RX MEDLINE=20573864; PubMed=11124544;
RA Tascou S., Uedelhoven J., Dikens C., Nayeria K., Engel W.,
RA Burfelnd P.;
RT "Isolation and characterization of a novel human gene, NF3L1, and its
RT mouse ortholog, Nif3l1, highly conserved from bacteria to mammals.";
RL Cytogenet. Cell Genet. 90:330-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21100893; PubMed=11161814;
RA Hadano S., Yanagisawa Y., Skaug J., Fichter K., Nasir J.,
RA Martindale D., Koop B.F., Scherer S.W., Nicholson D.W., Rouleau G.A.,
RA Ireda J.E., Hayden W.R.;
RT "Cloning and characterization of three novel genes, ALS2CR1, ALS2CR2,
RT and ALS2CR3, in the juvenile amyotrophic lateral sclerosis (ALS2)
RT critical region at chromosome 2q33-q34: candidate genes for ALS2.";
RL Genomics 71:200-213(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA Mao Y.M., Xie Y., Huang X.Y., Ying K., Dai J.L.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX Huang C., Qian B., Yu Y., Gu W., Wang Y., Han Z., Chen Z.;
RT "Novel genes expressed in hematopoietic stem/progenitor cells from
RT myelodysplastic syndromes patient.";

RP SEQUENCE FROM N.A.
RC TISSUE-Skin:
RX MEDLINE-22388257; PubMed-12477932;
RA STRAUSBERG R.L., Felngold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhut N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshlyak S., Carinici P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Maden A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Roderfield J.S.N., Krzywinski M.T., Skalska U., Smalins D.E.,
RA "Generation A., Schein J.E., Jones S.J.M., Marra M.A.:
RT "Severing and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF) FAMILY.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb.ch).

CC DR EMBL: AF283538; AAC44846.1; ALT_INIT.
DR EMBL: AB038949; BAB32499.1; -
DR EMBL: AF060513; AAG43131.1; -
DR EMBL: AF182416; AAG44952.1; -
DR EMBL: AK023378; BAB14551.1; -
DR EMBL: BC007654; AAH07654.1; ALT_INIT.
DR Genew: HGNC:13390; NIF3L.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3_1.
DR Trifams: TIGR00486; TIGR00486_1.
FT CONFLICT 77 82 TWNTWK -> NLEHNGR (IN REF. 4).
SQ SEQUENCE 350 AA: 38983 MW: 81FLASAD5JBZSD7 CRC64;

Alignment Scores:
Pred. NO.: 1.37e-143 Length: 350
Score: 1799.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.50% Indels: 0
DB: 1 Gaps: 0

245 ATGATTTGAAGCGCTCCTTCCTTGCAACTTGACATCCCTCGTTGTCTGAG 304
|||||

[illegible]

AC O9E080; Q9D098; 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NIF3-Like protein 1.
 GN NIF3L1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20573864; PubMed=1124544;
 RA Tascou S., Uedelhoven J., Dlxkens C., Nayerla K., Engel W.,
 RA Burelnd P.,
 RT "Isolation and characterization of a novel human gene, NIF3L1, and its
 RT mouse ortholog, Nif3l1, highly conserved from bacteria to mammals.",
 RL Cytogenet. Cell Genet. 90:330-336(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant P.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,
 RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Komboerts P.,
 RA Nozaki P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF284439; AAG45961.1; ALT_INIT.
 DR EMBL: AK011670; BAB27769.1; ALT_INIT.
 DR MGD: MGI:1929485; NIF3L1.
 DR InterPro: IPR002678; DUF34.
 DR Pfam: PF01784; NIF3; 1.
 DR TIGRfams: TIGR00486; 1.
 DR CONFLICT 164 164 G -> S (IN REF. 2).
 FT CONFLICT 178 178 E -> K (IN REF. 2).
 FT CONFLICT 191 191 L -> F (IN REF. 2).
 FT CONFLICT 195 195 L -> O (IN REF. 2).
 FT CONFLICT 198 198 L -> F (IN REF. 2).
 FT CONFLICT 208 208 T -> I (IN REF. 2).
 SQ SEQUENCE 350 AA; 38828 MW; B0FA71503FCF7086 CRC64;

Alignment Scores:
 Pred. No.: 1,03e-124 Length: 350
 Score: 1575.00 Matches: 302
 Percent Similarity: 93.14% Conservative: 24
 Best Local Similarity: 86.29% Mismatches: 24

Query Match: 56.47% Indels: 0
 DB: 1 Gaps: 0
 US-09-745-506-74 (1-1553) x NIF3L_MOUSE (1-350)
 QY 245 ATGATTTGAAGGCTCTCTTCTTCTGAATGATTTGCAATCCCTCTGTTGCTGAG 304
 DB 1 MetaPrLeuLylalaLeuLeuSerLeuLeuAnaPhaLalSerLeuSerPhaLagL 20
 QY 305 AGTTGGGACATGTTGGATTACTGTTGGAACCAAGCCACACATCTGTAATACATC 364
 DB 21 SerTPAspAsnValGlyLeuLeuValGluProSerProPhoHisThrValAsnThrLeu 40
 QY 365 TTCCTGACCAATGACCTGACAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
 DB 41 PhleuThrAsnAspLeuThrGluGluValMetAspGluLalaLeuLInysLysLalaAsp 60
 QY 425 CTCATCTCTCTTACCATCCGCTTCTCTCCAGCCATGACCGCATCTGACACACA 484
 DB 61 PhelleuSerThrHisProProlePhaPromeLysHisLethrTrpLysThr 80
 QY 485 TGGAGAGAGCGCTGATCCGGCTGTGAGACAGAGTGGATCTCTCTCTAT 544
 DB 81 TrpLysGluCylLeuValLleArgLalaLeuGluAsnArgValAlaValLysSerProHis 100
 QY 545 ACAGCCTATGATGCTGGCGCCGACGGCGCAACCACTGTTGGCTAAAGGCTTGGAGCT 604
 DB 101 ThrAlaTrpAspLalaLalProGluGluValAsnSerTrpLeuLalaLysLLeuGluLThr 120
 QY 605 TGTACCTCCAGGCCATACATCTCTCCAAAGCTCCCACTACCTACAGAGGAAACAC 664
 DB 121 CystThrThrArgProIleHisProSerArgLalaProAspTyProthrgLugLalaHis 140
 QY 665 CGATGATTTACAGTTAACTATACACCCCAAGACCTGGACAAAGTCTGCTGACAGGAA 724
 DB 141 ArgLeuGluPhaSerValAsnArgSerGluAsnAspLysValMetSerThrLeuArg 160
 QY 725 GGAATGAGCGGTTCTCTCACTCTTTCTGAGACAGCTGTAATAGGAACAACA 784
 DB 161 GlyValGlyLylValSerValThrSerPheProAlaArgCysAspLylGluGluLThr 180
 QY 785 CGGATTAATCTGAATGTACTCAGAAAGCTTTGATGAGAGTGTGATTTCTTCCCG 844
 DB 181 ArgLleSerLeuAsnGlyThrGluLysThrLeuMetGluValLeuLalPheLeuSerGlu 200
 QY 845 AACCAAACTTTATGAGAAGACGGAATTTCTACTGAGAAAGCTTTGCTTACT 904
 DB 201 AspArgGluLeuTylGluLysThrGluLleLeuSerLeuGluLylProLeuLeuHis 220
 QY 905 ACTGGAATGGAGCGGTTATGACACACTGATGATCTGTCCTGGACAACTGATGAT 964
 DB 221 ThrGlyMetGlyArgLeuGlyThrLeuAspGluSerValSerLeuLalaLleMetLleGlu 240
 QY 965 CGAATTAAGAAACACCTAAACTATCTATATCTGATCCCTTGGGGTGGGAGAAC 1024
 DB 241 ArgLleLysThrHisLeuLysLeuSerHisLeuArgLalaLleuLylGlyArgThr 260
 QY 1025 TTAGAGTCTCAAGTCAAAATCGTGGCGCTGTGCTGTCTTGGGAGACAGCTTGGCAG 1084
 DB 261 LeuGluSerGluValLylValAlaLalaLeuGlyAspLylSerLylLysSerValLleGlu 280
 QY 1085 GGGTGGAGCGTACCTTATACCTCACAGGAGATGATCCCATCATGATCTTGGATGCT 1144
 DB 281 GlyValGluAlaAspLeuThrLeuThrGlyGluMetSerHisHisAspValLleuAspAla 300
 QY 1145 GCTTCCCAAGGAATAATGTCACTCTGTGACACAGCAACACTGACAGAGGCTTCTT 1204
 DB 301 AlaSerLysGlyLleAsnValLleLeuGlyGluHisSerAspThrGluArgLylPheLeu 320
 QY 1205 TCTGACCTTCGAGATTCGTGATTTCTCACTTGGAGAAATGAATATATATCTCTATCA 1264
 DB 321 SerGluLeuGluGluMetLeuGluLylHisPheGluAsnLysLleAsnLleLleuSer 340

1265 GAGACGTACAGGAGGACCCCTTCGAGGCGTA 1294
 |||||||||||||||||||:|||||
 341 G|U|T|H|A|S|P|A|G|A|S|P|R|O|T|E|A|L|Y|A|L 350

RESULT 3
 A36A_DROME STANDARD; PRT; 292 AA.

AC Q9NKK5; O9VJL4; 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Anon-35F/36A protein.
 GN ANON-35F/36A OR BG:DS02740.16 OR CG4278.
 OS Drosophila melanogaster (Fruit Fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10471707;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wotman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blatz R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abdl J.F., Asgari A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Baas A.A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Butlis K.C., Busam D.A., Butler H., Caddieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.P., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Moulton G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclow J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinelt K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.E., Spradling A.C., Stapleton M., Stokrup L., Sun E.,
 RA Svirydas R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri N., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).

CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.

```
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-slb.ch/announce/ or send an email to license@isb-slb.ch).
```

```
DR EMBL; AE003415; AAF65006.1; -  
DR EMBL; AE003650; AAF653525.1; ALT_INIT.  
DR Flybase; FBgn0014093; CG4278.  
DR InterPro; IPR002678; DUF34.  
DR Pfam; PF01784; NTF3_1  
DR TIGRfams; TIGR00486; TIGR00486; 1  
SQ SEQUENCE 292 AA; 32363 MW; 775D2E7721150CC CRC64;  
  
Alignment Scores:  
Pred. No.: 1,01e+40 Length: 292  
Score: 578,00 Matches: 118  
Percent Similarity: 51,44% Conservative: 61  
Best Local Similarity: 33,91% Mismatches: 81  
Query Match: 20,72% Indels: 88  
DB: 1 Gaps: 4  
  
US-09-745-506-74 (1-1553) x A36A_DROME (1-292)  
  
QY 251 TTGAAGGCTCCTCTTCCTTGTAATGACTTTGCATCCCTCGTTCGTGAGATTGG 310  
Db LeuLaLaLaValVallylSgLuLeuLunspPheAlaProThrSerTrpAlaIleuylStrp 32  
QY 311 GACAACTGTGGCATTAAGTCTGGTGGAACCAAGCCACACACTACTGTAATAATCACTCTTCG 370  
Db AspaSnValGIyLeuLeuIleGlurProHisArgIuLySgluIleLylysIleLeuLeu 52  
QY 371 ACCAATGACCTGACTGAGAAGTAGATGGAGAGGTGCTGCAAAAAGAGGACAGACTCAT 430  
Db ThrsnAspLeuThrInglurProValIallylSgluAlaLeuGIuLySgluAlaIleuile 72  
QY 431 CTCCTCAACATCCGCCCTATCTTCCGACCATGAAGCGCATTAACCTGGAACACATGANG 490  
Db IleSerTyHISproProIlePhelysProLeuThrArgIleThrgIasnerHisTrypyls 92  
QY 491 GAGGGCCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACCTCCATCACAGCC 550  
Db GluUrArgValValaLaLaLaCyseLeuLaInsnpIleAlaLeuTySerProHisThrAla 112  
QY 551 TATGATGTGGGCCCCCAGCGGCTCAACAACACTGGCTGGCTAAAGGCGCTTGAGACTGTACC 610  
Db TrpsAsplyLySserGIyValaInsnsprtleuSerLySaLaValaInsIleIleSer 132  
QY 611 TCACAGGCCCATATACCTCTTCCAAAGCTCCCAACTACCTACAGAGGAGAAACACCGAGTA 670  
Db IleHrgrProLeuGIuPrOglu----- 139  
QY 671 GAATTCACACGTTAACTACACCCAGACCTGGACAAAGTCATGTCTGCAGTGAAGAAT 730  
Db LeuGIyAlaPrObrogly---- 145  
QY 731 GACGGTGTTCGTGACTCTTTTTCTGCTAGAGACTGGTAATGAGGAACAACAGGANT 790  
Db ----- 145  
QY 791 AATCGAATTTGTAACAGAGGCTTTGATGCAGGTGATGATTTCTTCCGGAACANA 850  
Db ----- 145  
QY 851 CAACCTTATCAGAAAGCGAAATTTGTCTCACTGGAAGAACCTTTCCTTACATACCTGA 910  
Db -----Thrgly 147  
QY 911 ATGGGACGGTATATGACACTGGATGAATCTGCTCCCTGGCAACACATGATGATGAANA 970  
Db ----- 148 serGIyAlgyTyr-----IleGIuThrLySmetGIuLeuSerGIuValaJalInsSerLeu 165
```



```

QY 953 ACCGATGATGATGCAATAAAAAGACACCTAAACATATCTATCTGCTTACGACCTGGG 1011
Db 163 GtUeuValGlnAldgAlaLysGtUeuThcGlyUeuGlnTyrValGlnValCysAlaPro 182
QY 1013 GTGGGGAGAACCTTAGAGTCTCAAGTCACAACTGCGGCGGCTGTGTGCTGTGGAGC 1072
Db 183 AsnGly-----LeuAspSerHisIleSerLysValSerLeuCysAlaGlySerGly 200
QY 1073 AGCGTCTGCAGGGGTGTGAGGCGTCACTTAACTCACTCAAGGTGAGAGTGCACATGAT 1132
Db 201 SerValValMetAsnThrAspAlaAspLeuTyrThrGlyGtUeuSerHisGln 220
QY 1133 ACTTGGATGCGCTGCTCCAGGAAGATAAATGCTATCTGTGAGACAGACACATGAA 1192
Db 221 ValUeuAlaMetAlaMetAlaLysGlyLleSerValLleUeuCysGlnHisSerAsnThrGlu 240
QY 1193 CGAGCGCTTCTTCTTGACCT-----CGAGATATGCTGATTTCT---CACTTGCAGAT 1243
Db 241 ArgGlyLysTyrLeuLysAspValMetCysGlnLysLeuAlaSerSerHisLysGly 260
QY 1244 ---AAGATAAATATATATCTATCTCAGAGACAGACAGGACCGCTTT 1285
Db 261 ValAspAlaAsnValLleValIleSerSerMetAspAlaAspProLeu 275

RESULT 5
NIF3_YEAST STANDARD; PRT; 288 AA.
ID NIF3_YEAST
AC P53081;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NGC1-interacting factor 3.
GN NIF3 OR YGL221C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
[1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97435481; PubMed=9290212;
RT Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.";
RL Yeast 13:1077-1090(1997).
-1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use by
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC CC EMBL; 272743; CAA96937.1; -.
CC DR PIR; S64243; S64243.
CC DR SGD; S0003189; NIF3.
CC DR InterPro; IPR002678; DUF34.
CC DR Pfam; PF01784; NIF3; 1.
CC DR TIGFAMS; TIGR00486; TIGR00486; 1.
CC SO SEQUENCE 288 AA; 31888 MW; B6AB6E48AFA776A CRC64;

Alignment Scores:
Pred. No.: 2, 03e-24 Length: 288
Score: 384.50 Matches: 104
Percent Similarity: 44.78% Conservative: 98
Best Local Similarity: 28.57% Mismatches: 59
Query Match: 13.79% Indels: 103
Gaps: 11

US-09-745-506-74 (1-1553) x NIF3_YEAST (1-288)

```

QY	251	TTGAAGCCTCTCCCTTTCTCTTGAAAGATTTGATCGATCCCTCGTTTGGTGAAG---AGT	307
Db	10	LeuAspLysLeuValaIarGserIleThrLysPheYrProIInLysTryAlaAspLysSer	29
QY	308	TGGGACAATGTGGATTTACTGTGTGGACCAAGCCACACATCTGTAATATCA-----	361
Db	30	TrpAspSnnThrgLLeuLeuIleAspCysSerThzAlaGlnValThrThrAlaAspAla	49
QY	362	-----CTCTTCCTGACCAATGACCTGACTGAGGAAGATGAGGAGAGCTG	406
Db	50	AsnAlaLysThrLysValLeuLeuThrValAspLeuThrLysSerValaIaGlnGlnAla	69
QY	407	CTGCAAAAAGAGCAGACACTCATTTCTCTCTACACCTCGGCTATCTTCCGACCCATGAG	466
Db	70	ValAspAlaAsnLysAsnValIleValaIaLysTrisProhellePheProSerTrpAsn	89
QY	467	CGCATTAACCTGGAAACA-----TGAAGAAGACCGCTGGTGAATCGGGCTGGGAACAAGA	523
Db	90	ArgLeuSerProhIsthrAsnProGlnIleGluThralaIleLysLeuIleGlnTryGly	109
QY	524	GTCGGATTTACTCTCTCTCTATACAGCCTATGATGCTCGGCCACGAGCGCTCAACATCGG	583
Db	110	IleSerValTyrCysProhIsthrAlaValaIaAspAlaIaArgLysGlyValaIaAsnAspTrp	129
QY	584	TTGGCTTAAGGCGCTTGGAGCTTGTACTCCAGCGCCATACACTCTTCCAAAGCTCCAC	643
Db	130	LeuValaIaArgLysLeuAsn-----	135
QY	644	TACCTCTACAGAGGAAACCAACCGAGTAGAATTCAAGCTTAACCTACACCCAGACCTGGAC	703
Db	135	-----	135
QY	704	AAAGTCATGTCTGCAGTGAAGAAGATTTAGCGGTCTTCTGTCACT---TCTTTTTCGCT	760
Db	136	-----AsnIleGlnAsnValaIaLysSerTryAlaLeu	146
QY	761	AGGACTGGTAATGAGGAACAACAGCGATTATCTGAATTGTACTCAGAAAGGCTTGTATG	820
Db	147	GluThrValSerGlyLeu-----	152
QY	821	CAGGTGGTAGATTTCTTTCCCGGACAAACAACTTATCAGAAGACGGAAATTCTGTCA	880
Db	153	-----ThrAspAspLeu---	156
QY	881	CTGAGAAAGCCTTTGCTTTCTACATACCTGAGATGGAGCGTTATGCACACTGGATGAATCT	940
Db	157	-----IleGlyTyrGlnArgPheValGluPheAsnLysAsp	168
QY	941	GTCCTCCCTGCGAACCATGATTTGATCGAATAAAGACACSTAAACATATCTCATATTGCG	1000
Db	169	IleSerLeuGlnIleValLysAsnValLysArgValIleAlaArgValProTyrValGln	188
QY	1001	TTAGCCCTTGGGGGTGGGGAACCTTAGAGTCTCAAGTCAAAAGTC-----GTGGCCCTG	1054
Db	189	ValaIaSerLeuAlaIaIaProSerAlaTrpAsnGlnLeuLysIleLysValaIaIaIaI	208
QY	1055	TGTACTGGTGTGGGAGCAGCGTCTCGACGGGTGT-----GAGGCGACCTTTTACCT	1108
Db	209	CysAlaIeLysSerGlySerGlyValaPheLysGlnLeuLysGlnAspValaIaAspLeuTyrTr	228
QY	1109	ACAGGTGAGATGCCCATCATGATTACTTTGGATGCTGCTTCCCAAGAAATAATGTCAATC	1168
Db	229	ThrgLysIleMetSerIleAsnIleGlnValLeuLysTrpLysGlnMetGlnLysThrValIle	248
QY	1169	CTCTGTGAACACAGCAACAGTACAGAGCGAGCTTTCTTCTGACCTTCCGAGATGATGCTG---	1223
Db	249	ValCysAsnIleSerAsnThrgLnaIaArgLysPheLeuGlnAspValaIaMetLysGlyLeuLeu	268
QY	1226	-----GATTCCTCACTTGGAGAAATAAGATAATATATCATCTATCAGAGACTGACAGGAGC	1279
Db	269	GlnAspGlnGlnLysIle-----GluValaIaValSerLysMetAspCysAsp	283
QY	1280	CCTCTTCAGGTG	1291


```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessleres P., Bolotin A., Borchert S.,
RA Borrias R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.F., Connerly I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fader C., Ferrazi E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Gutserpi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsch S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurlita K., Lapides A., Lardinois S., Lauber J., Lazarovic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rhyolita C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowaka A., Seror S.J., Serron P., Shin B.S., Solido B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Terasaki M., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Wetzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yamamoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1 SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -1 CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 270.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D84432; BAA12492.1; ALT_FRAME.
DR EMBL: 299116; CAB14447.1; -.
DR PIR: A69954; A69954.
DR Subtilist: B61161; YqfO.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3.1.
DR TIGRFAMS: TIGR00486; TIGR00486.1.
KW Hypothetical protein: Complete proteome.
FT CONFLICT 265 266 KD->FC (in REF. 1).
SQ SEQUENCE 373 AA; 40920 MW; E2320C1268DBE9B CRC64;
Alignment Scores:
Pred. No.: 4.9e-16 Length: 373
Score: 285.50 Matches: 99
Percent Similarity: 46.84% Conservative: 79
Best Local Similarity: 26.05% Mismatches: 143
Query Match: 10.24% Indels: 59
DB: 1 Gaps: 17
US-09-745-506-74 (1-1553) x YQFO_BACSU (1-373)
QY 260 CTCCTTCTCTTAATGACTTTCATCCCTCGTGTGCTGAGAGTGGACATGTT 319
DB 10 lletleaglneuphlegluphserProlysalaItyralaValaGluGlyspysile 29
QY 320 GGATTA-----CTGGTGAGACCAAGCCACACATACGTGTAATGACTCTTC 367

```

```

DB 30 GlyLeuGlnIleGlyThrLeuAsnLysPro-----IleLysAsnValMet 44
QY 368 CTGACCAATGACCTGACTGAGGAAGTATGAGAGGATGTCGAAAGAGGACACCTC 427
DB 45 ValThrLeuAspValLeuGlnSerValIleAspGluAlaIleGluGlyValAspLeu 64
QY 428 ATTCTCTCTTACCATCCGCTTATCTCCGACCATGAGCCGATTAACCTGGACATG 487
DB 65 lletlealIahIshIspProIlePheArgSerLeuLysIshIleSerThrAspIlePro 84
QY 488 AAGAGGCCCTGGATGCTGGGCTGTGGAGAACAGAGCGGATCTGCTCATCA 547
DB 85 AlaGlyArgLeuIleGluLysCysLeuLysIshAspIleAlaValTyraIahIshThr 104
QY 548 GCCTATGATGCTGCGCCCGAGGCTCAACACGTTGGCTTAAGGCGTGGACCTGT 607
DB 105 AsnLeuAspValAlaAspGlyValAsnAspLeuAlaGluAlaLeuGlnLysSer 124
QY 608 ACCTCCAGGCCATCATCTTCCAAAGCTCCCACTAC 646
DB 125 GluThrGluValLeu-----AlaProThrTyThrAspProLeuLysLysLeu 140
QY 647 -----CCTACAGAG-----GGAAC----- 661
DB 141 AlaValTyThrValProLysGluTyThrGluGlnValAlaGluAlaLeuGlnLysAlaGly 160
QY 662 -----CACCGATGAATTCACGTTACCTACACCCAGACCTGGACAACTGCT 715
DB 161 AlaGlyIshIleGlyGluTyThr-----SerIshCysAlaPheSerSerGluLysIleGly 178
QY 716 GCAGTGAAGAGATGAGGCTGTTCTGCTACTTCTTTCGTCAGAGCTGTAATGAG 775
DB 179 SerPheLysProLeuAspValAlaLys-----ProPheIleGlyGluValLysLysLeu 196
QY 776 GAA-----CAACAGCGATTAATCTGAATTTACTACAGAGCTTGTGATGAGCTG 826
DB 197 GluLeuValIshGluValAlaArgLeuGlnThrValPheProLysSerValGluLysAlaVal 216
QY 827 GTGATTTTCTTTCCCGGAAACAACTTATTCAGAAAGC-----GAAATTGTGCA 880
DB 217 lletasnaIameIleLysSerIshPro-----TyrgluGluValAlaTyraLysPleTyThrPro 235
QY 881 CTGAGAGAACCTTCTGCTTCTACATCTGGAAGGAGCGTTTGTGACACTGATGAATCT 940
DB 236 ValGluGln--ThrProAlaGluLysGlyLeuGlyValGlyValGlyThrLeuLysAsnGlu 254
QY 941 GCTCCCTCGGCAACCATGATGATGAATTAATAAAGACACTTAATCTCATATTGCTGC 1000
DB 255 MetThrLeuLysGluPheAlaLeuPheValLysAspLysLeuAspValAsnGlyValArg 274
QY 1001 TTAGCCCTTGGGGTGGGAGAACCTTAGAGTCTCAAGTCAAGTCAAGTGGCCCTGTGCT 1060
DB 275 Met-----ValGly---AspAlaSerSerMetValLysValAlaValLeuGly 290
QY 1061 GGTTCGGAGACGAGGCTTCGAGGCTGTGAG-----GCTGACCTTCACTCA 1111
DB 291 GlyAspGlyAsnLysTyThrIshIshAlaLysAspGlyGlyValAspValTyraIshThr 310
QY 1112 GGTGAGATGCTCCCATGATGATCTTGTGATGCTGCCCAAGGATTAATGCTCATCTC 1171
DB 311 GlyAspLeuTyThrPheIshValAlaIshAspAlaMetMetLeuGlyLeuAsnValTyraLysP 330
QY 1172 TGTGAACAGCAACAGCACTGAACGAGGCTTCTTCTGACATCGAGATGCTGATCT 1231
DB 331 ProGlyHis--TyraLagluLysIleMetLysGluGlyValThrArgLysLeuThrSer 349
QY 1232 CACTTGAGATTAAC-----ATTAATATTATTCATTCAGACACTGACGAGACCT 1282
DB 350 MetCysAsnAspLysLysPheGlyValAsnIlePheValSerGluThrAspThrAsnPro 369
RESULT 8
YF59_STAMM

```


ID YF59_STAMM STANDARD: PRT: 366 AA.
AC 09977;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SAV159/SAI388.
GN SAV159 OR SAI388.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158678, 158679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsunari H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekitani K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.
RL Lancet 357:1225-1240(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF003362; BAB57721.1; -
DR EMBL: AF003134; BAB42651.1; -
DR PIR: F89936; F89936.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3; 1.
DR TIGRPFAM: TIGR00486; TIGR00486; 1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 366 AA; 41075 MW; CA7D5FADF78FFFA CRC64;

Alignment Scores:
Pred. No.: 2 09e-15 Length: 366
Score: 278.00 Matches: 105
Percent Similarity: 42.57% Conservative: 67
Best Local Similarity: 25.99% Mismatches: 140
Query Match: 9.97% Indels: 92
DB: 1 Gaps: 18

US-09-745-506-74 (1-1553) x YF59_STAMM (1-366)
QY 245 ATGGATTGAAAGGCTCTCTTCCTTCGAAATGATTCATCCCTCTGTTGTGAG 304
DB 1 MetLysIleAlaSpruMetThrLeuLeuAspHisHisValProPheSerThrAlaGlu 20
QY 305 AGTTGGACAAATTTGATTACTGCTGGAACCAAGCCACCATATCTGTAATACACTC 364
DB 21 SerTrpAspAsnValIleLeuLeuIleGlyAspValAlaGluValThrGlyValLeu 40
QY 365 TTCCCTGACCAATGACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 424
DB 41 ---ThrIleLeuAspCysThrLeuGluValAlaAsnGluAlaIleGlyValGlyTyrAsn 59
QY 425 CTCATTCTCTACATCCGCTATCTTCGACCAAGGACATGAAGGACATTAACCTGGAACA 484
DB 60 ThrIleIleSerHisProLeuIlePheLeuGlyValThrSerLeuValAlaAsnGly 79
QY 485 TGAAGAGCGCTGATCGCGGCTCTG---GAGACAGAGTGGTATCTACTCTCT 541
DB 80 Tyr-----GlyLeuIleIleArgLysLeuIleGlnHisAspIleAsnIleAlaMet 97

QY 542 CATACAGCCATGATGCTGGCCCAAGGCGCTCAACAATGTTGGCTAAAGCGTTGA 601
DB 98 HisThrAsnLeuAspValAsnProTyrGlyValAlaAsnMetLeuAlaValAlaMetGly 117
QY 602 GCTTGTACCTCCAGGCCCATACATCTTCCAAAGCTCCACATACCTTACA----- 652
DB 118 LeuLysAsnIleSerIleIleAsnAsnGlnGlnAspValTyrTyrValGlnThrTyr 137
QY 652 ----- 652
DB 138 IleProLysAspAsnValGlyProPheLysAspLysLeuSerGluAsnGlyLeuAlaGln 157
QY 653 GAGGGAACACCGAGTAGAATTC----- 676
DB 158 GluGlyAsnTyrGluTyrCysPhePheGluSerGluGlyArgGlyGlnPheLysProVal 177
QY 677 ---ACGTTTACTACACAC---CAAGACCTGACCAAGTCAATGTCGACGTGAAGAAT 730
DB 178 GlyGluAlaAsnProThrIleGlyGlnIleAspLys-----IleGluAspVal 193
QY 731 GACGGTGTCTTCTGCTACTTCTTTTCTGCTAGCAGCTGGTAATGAGAACACAGCAT 790
DB 194 AspGluValLysIle-----GluPhe 200
QY 791 AATCTGAATGTACTCAGAGCGTTGATGACAGGTGATATTTCTTCCGGAACAA 850
DB 201 MetIleAspAlaTyrGlnLys-----SerAlaGln 211
QY 851 CAACCTTATCAGAG-----ACGAAATCTGTCACTGAG 886
DB 212 GlnLeuIleLysGlnTyrHisProTyrGluThrProValPheAspPheIleGlyLeuLys 231
QY 887 AAGCCTTGTCTTACTACTGCAATGGAGCGTTATGACACACGTGATGATCTGTCC 946
DB 232 GlnThrSerLeuTyr-----GlyLeuGlyValMetAlaGluValAspAsnGlnMetThr 249
QY 947 CTGGCAACATGATTCATGCAATAAAGACACCTAAACTATCTATTCGCTTACGCC 1006
DB 250 LeuGluAspPheAlaAlaAspIleLysSerLysLeuAsnIleProSerValArgPhe--- 268
QY 1007 CTGGGGTGGGAGAACCTTAGAGTCAAGTCAAAAGTCTGGCCCTGTGTCTGTCT 1066
DB 269 -----ValGlyGluSer---AsnGlnIleLysArgIleAlaIleIleGlyGlySer 285
QY 1067 GGG-----ACAGCGTCTCGAGGCTGTGAGCTGACCTTACTACATA 1111
DB 286 GlyIleGlyTyrGluTyrGlnAlaValGlnGlnGly-----AlaAspValPheValThr 303
QY 1112 GGTGAGATGCCCATCATGATTCATTCGATTCGCTTCCCAAGGAAATAATGTCATCTC 1171
DB 304 GlyAspIleLysHisHisAspAlaLeuAspAlaLysIleHisGlyValAsnLeuIleAsp 323
QY 1172 TGTGACACAGCAACACTGAAGAGGCTTCTTTCACCTTCAGATATGCTGAGTCT 1231
DB 324 IleAsnHis---TyrSerGluTyrValMetLysGluGlyLeuLysThrLeuMetAsn 342
QY 1232 CACTTGGAG---AATAGATTAATATATCTCTA-----TCAGAGATGACAGGACCT 1282
DB 343 TrpPheAsnIleGluLysIleAsnIleAspValGluAlaSerThrIleAsnThrAspPro 362
QY 1283 CTTCAGCGTGTGA 1294
DB 363 PheGlnTyrIle 366

RESULT 9
ID YD03_CIOAB STANDARD: PRT: 268 AA.
AC 09710;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein CAC1303.

or send an email to license@lsb-sib.ch.

CC EMBL; AB001511, BAB05099.1, -
 DR PIR; D83822, D83822.
 DR InterPro; IPR002678, DUF34.
 DR Pfam; PF01784, NIF3.1.
 DR TrEMBL; TIGR00486, TIGR00486; 1.
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 372 AA; 40813 MW; DD79BF630DCDE5B CRC64;

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Query Match:	Gaps:
3.5e-14	372	263.50	89	80	24.938	141	47	9.458	12

US-09-745-506-74 (1-1553) x YD80_BACHD (1-372)

QY 317 GTTGAATTAAGTGGTGGACCAAGCCACACATACATGTAATACATCTTCCTGACCAAT 376
 DB 32 IIEGLYTHRLHLSNLSYPRO-----IIEGLNRYVALLHUTHRALALEU 46
 QY 377 GACCTGACTGAGAACTGATGAGAGAGTCTGCACAAAGAGCAGACCTCATCTCTCC 436
 DB 47 AARVALTHRGUSERVALILEASRGVALLIEGLULEGLYLAIEGLULEUALLIEUALLA 66
 QY 437 TACCAATCCGCTATCTCCGACCATGAAAGCCATACCTGACACACATGGAAGAGCC 496
 DB 67 HISHASPRGIIIELEHETARGPROLEUSERSERILEARGTHRSAPRHLATYRGLYATG 86
 QY 497 CTGCTGATCCGCGCTGGACACAGATCGATACCTGATACCTCATACAGCTTATGAT 556
 DB 87 IIEGLULYALALEYSHASPRLEUTHRIETRYALALAHSTHANSLEUASP 106
 QY 557 GCTGCGCCCGAGGCGCTCAACAATGTTGGCTAAAGGCGTTGACCTGCTCAGG 616
 DB 107 ILETHLYGVLGYLVALYANASPRLEUMETALASRALALEUGLYLEUYSAPRIEGLU 126
 QY 617 CCATATACCTCTCCAAAGCTCCCACTACCTACAGAGGAAACCCAGAGTGAATTC 676
 DB 127 VALLEALAPROTHR-----THTHGLISERLEUUYRYSLEUVALVAL 141
 QY 677 AACGTAACTACACCAAGACCTGACAAAGATCATGCTGCTGACGAAAGGATTCAGGT 726
 DB 142 RHEVALPRGHIHSTHRLIS---THRASRGVALLARGVALLALEUGLYATGVALA 160
 QY 737 GTTCTGCTACTCTTTCTGCTAGGACTGTAATGAGGACAAACA----- 784
 DB 161 GLYHLSIIEGLYANATYRSEYRYSYTHRPHESANSERLYSGLYTHRGYLYTHRPHLYS 180
 QY 785 -----CGGATTAATCTGATTTGATCTACAGAAAGCTTTGATGACAGTGGATTTT 835
 DB 181 PROGLUGLYLYTHRSNPRORHEIIEGLYLYSGVGLYALALEUGLYRHEVALIGLU 200
 QY 836 STT-----TCCCGAACCAACAACCTTATTCAGAGACG 868
 DB 201 LEULYIIEGLULHRLHLEVALTHRGVGLYGLYLSNLSYVALYALALEALAMELIE 220
 QY 869 GAATTTCTGCTACGAGAAAGCTTTGCT--CTACAT----- 904
 DB 221 LYSERHNPRTGTYLUGLUPROALATYRASPRLTYRTOLEUALLANSGLUGLU 240
 QY 905 ---ACSGAAATGGAGGATTTATGACACATGATGATCTCTCCCTGGGACCAATGAT 961
 DB 241 THLEUGLYLEUVALYATGIIIEGLYLYTHLHNSISGLUSERMETHRLHNSGLUGLUPHEA 260
 QY 962 GATCGAATATAAAGACSTAAACATATGATCATATGCTTACCCCTGGGAGGAGAGA 1021
 DB 261 LYSGLVVALYLYSVALAPHEASPRVALPYOTHLALARGVAL-----VALGLY--- 276
 QY 1022 ACCTTAGAGTCTCAAGTCAAAAGTGTGGCCCTGTGCTGCTGCTGAGGACAGCTTCTG 1081

DB 277 SERLEUGLHTRGHLNLEARGLYSVALALAVALLEUGLYLYSPRGYANLSTYRMET 296
 QY 1082 CAGCGTGT-----GAGCTGACCTTTACCTACAGGTGAGATGTCATCATGAT 1132
 DB 297 ALAHLSALALEUARGLYSGVLYALASPRVALILEVALTHRGVLYASPRVALTYRHLISVAL 316
 QY 1133 ACTTTGGATGCTGCTCCCAAGAAATTAATGCTATCTCTGTGACACACACACTGAA 1192
 DB 317 PROHISAPRALAEUMETASPRGLYLEUANSIIEVALASPRGOLYHIS---ASNAVGLU 335
 QY 1193 CGAGCTTTCTTCTGACCTTCGACATATGCTGATTCATCTGAGAAATAG----- 1246
 DB 336 LYSIIELEMETYSLNGLYVALYLYSGVLYLEUGLYLYSLEUUEUYSAPRYLYSTYR 355
 QY 1247 ---ATTAATATTAATCTCATACAGACTGACAGGAGCCCTTTCAGGTGTA 1294
 DB 356 ASPTRHGLVVALALASERSEVALIHSTHASPRTOPRHEHRLHLE 372

RESULT 11

YE89_LISIN
 ID YE89_LISIN STANDARD; PRT; 373 AA.
 AC Q92B08;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 GN L1N1489.
 OS Bacteria.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.2.A.
 RC SRRAIN-CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Doman E., Dominguez-Bernal G., Duchaud E., Duratt L., Dussurget O.,
 Ertlan K.-D., Fsih H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuuprat G.,
 Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Roland J.-A., Voss H., Wehland J., Cossart P.,
 RT Comparative genomics of Listeria species.
 RL Science 294:849-852(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

DR EMBL; AL596168, CAC96720.1, -
 DR PIR; AH1618, AH1618.
 DR ListList; L1N01489, -
 DR InterPro; IPR002678, DUF34.
 DR Pfam; PF01784, NIF3.1.
 DR TrEMBL; TIGR00486, TIGR00486; 1.
 KM Hypothetical protein; Complete proteome.
 SO SEQUENCE 373 AA; 41436 MW; 4BBCA93B0B9F540 CRC64;

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:
1.5e-13	373	256.00	86	78	22.518	152	66

Query Match: 9.188

```

DB: 1 Gaps: 11
US-09-745-506-74 (1-1553) x YEB8_LISIN (1-373)
QY 272 TTGAATGACTTTGATCCCTCTCTGCTTGGATGGACATGTTGATCTACTGTC 331
DB 13 MetLysIleAlaProLysLysLeuAlaMetGluLysProIleGlyLeuGlnAla 32
QY 332 GAACCAAGCCCGACATACCTAGTAAATACACTCTTCGACCAATGACTGAGGAA 391
DB 33 GlySerLeuSerLysLysValArgLysValMetPhe---ThrLeuAspValLeuGlu 51
QY 392 GTGATGAGAGAGGTCTCCAAAGAGCAGACCTATCTCTCTACCATCCGCTATC 451
DB 52 ValValAspGlnAlaIleGluLysArgValAspLeuIleIleAlaHisIleProPheLeu 71
QY 452 TTGGACCCCATGAGCGCATACCTGAGACACATGAGAGAGCCCTGTATCCGGCT 511
DB 72 TyrArgProThrGlnHisIleAspThrThrThrLysGlnGlyLysMetIleLysLysLeu 91
QY 512 CTGGAGAACAGAGTCGATCTACTCTCTCATCATGAGCCTATGATGCTGCGCCAGGAGC 571
DB 92 IleLysHisAspIleThrValIleAlaHisIleThrAsnLeuAspIleAlaGlnGly 111
QY 572 GTCAACAACAGTGGTGAAGGCTTGAGCTGTATCCAGCCCATCATCTCTCC 631
DB 112 ValAsnAspIleLeuAlaAspLeuHisIleGlnAspThrThrMetIleGluGluThr 131
QY 632 -----AAGCTCCCACTACCTCCACAGG----- 655
DB 132 TyrThrGluProTyrCysLysIleAlaValIleValProGluAsnGluLeuGluSerVal 151
QY 656 -----GGAAACACACGAGTAGAATTCACGTTATACACCCAGACACCTG 700
DB 152 ArgLeuAlaLeuValAsnAsnGlyAlaGlyGlnIleGlyThrAsnTyrThrGlu----- 169
QY 701 GACAAAGTCATGCTGCAGTCAAAAGAAATGACGCTGTTCTGCTACTCTCTTCTCTCT 760
DB 170 -----CysThrPheHisIleThrGlyLysIleGlySerPheLysProGlyThrAspAlaAsn 187
QY 761 AGGACTGGTATGAGAGAACAAACGAGTATATCG----- 796
DB 188 ProThrIleGlyLysGluThrLeuThrSerIleProGluValLysIleGluAlaIle 207
QY 797 -----AATGTACTCAGAGGCTTGTATG----- 820
DB 208 PheProGlnTyrLeuThrGluThrIleThrLysAlaValLysIleAlaHisProTyrGlu 227
QY 821 ---CAGGTCGATGATTTCTTCCCGGAAACAAACTTATCAGAAACGGAATCTG 877
DB 228 GluProAlaIleAspValTyrThrLeuGluThrGlnThrTyrLysGlu----- 243
QY 878 TCACTGAGAGAGCTTCTCTACATCTGGAATGGAGGATTTGACACATGATGAA 937
DB 244 -----GlyLeuGlyAlaGlyValGlyIleThrLeuProLys 253
QY 938 TCTGTCTCTCGGACACCATGATGATGATTAATAAGACCTAAACCTATCTCATAT 997
DB 254 LysIleSerMetValSerPheIleAspLysLeuLysThrAlaPheAlaIleAspAsnVal 273
QY 998 CGCTTAGCCCTTGGGGTGGGAGAACCTTAGAGTCTCAAGTCAAAGTGTGCGCCTGT 1057
DB 274 ArgPhe-----ValGly---AspLeuLysAlaAsnValGlnLysValAlaIleIle 289
QY 1058 GCTGTTCTCGGAGAGCGCTCTG-----CAGGCTGTGAGCTGACCTTATACCTC 1108
DB 290 GlyIleLysAspGlyAsnLysPheIleHisGlnAlaLysAlaThrGlyAlaAspValPheIle 309
QY 1109 ACAGGTGAGATGCCCATCATGATTAATTTGATGCTGCTCCCAAGATTAATGTCATC 1168
DB 310 ThrGlyAspValTyrTyrHisIleThrAlaHisAspLeuAlaIleAsnLeuProThrIle 329
QY 1169 CTCGTGTGACACAGC-----AACACTGACAGAGGCTTCTTCTGACCTTCGACATATG 1222

```

```

DB 330 AspaIaGlyHisAsnIleGluLysValMetLysGlyTyrLeuLysAsnLysMetGlu 349
QY 1223 CTGATTTCTCAGCTTGAGACATAGATTAATTTATCCATGACAGACAGGAGCCCT 1282
DB 350 GlnAlaLysIleLeuAspTyrGlnAlaGluPheIleValSerGluValAsnThrAspPro 369
QY 1283 CTTCAG 1288
DB 370 PheGln 371
RESULT 12
YES2_LISMO STANDARD: PRT; 373 AA.
AC P53434;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein lmo1452.
GN lmo1452.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBL_Taxid=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO28 / Serovar 1/2c;
RC Klersfeld A.D., Cossart P.;
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ESD-e / Serovar 1/2a;
RC MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
RA Baquero F., Berche P., Blocher H., Brandt P., Chakraborty T.,
RA Charbit A., Cheouni F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Eutnan K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurepakt G.,
RA Nordeno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,
RA Nordstok G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
CC -!- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DB EMBL: U17284; AAA62502.1; -
DB EMBL: AL591979; CAC99530.1; -
DB PIR: AD1256; AD1256.
DB Listlist: lmo01452; -
DB InterPro: IPR002678; DUF34.
DB Pfam: PF01784; NIF3; 1.
DB TIGRFAMS: TIGR00486; TIGR00486; 1.
DB KW Hypothetical protein; Complete proteome.
DB SEQUENCE 373 AA; 41433 MW; C79AF8A87E2369A2 CRC64;
Alignment Scores:
Pred. No.: 3 35e-12 Length: 373
Score: 240.00 Matches: 88
Percent Similarity: 45.97% Conservative: 83
Best Local Similarity: 23.66% Mismatches: 155
Query Match: 8-61# Indels: 46
DB: 1 Gaps: 14

```

```

US-09-745-506-74 (1-1553) x YE52_LISMO (1-373)
OY 272 TTGAATGACTTTGATCCCTCGTTTGTGAGAGTGGACAAATGTGATTACTGNG 331
Db 13 MetGluLysIleAlaProLysLysLeuAlaMetGluLysProLysLeuGluVal 32
OY 332 GAACCAAGCCCAACATACATGTAATATACATCTTCCATACCAATGCTTACTGAGAA 391
Db 33 ---GlyAspLeuSerArgLysValArgLysIleMetPheThrLeuAspValLeuGlu 51
OY 392 GTGATGGAGAGAGTGTGCAAAAGAGACCTTATCTCTCCATACACCGCTTATC 451
Db 52 ValValAspGluAlaIleLeuLysValAspLeuIleLeuAlaHisPropLeu 71
OY 452 TTCGACCCATGAGGCAATACCTGACACATGAGAGAGAGCGCTGGATCCGGGCT 511
Db 72 TyrArgProThrGlnHisIleAspThrThrLysGlnGlyLysMetIleLysLeu 91
OY 512 CTGAGAACAGATGCGGTATCTACTCTCTCATACAGCTATGATGCTGCGCCAGGC 571
Db 92 IleLysHisAspIleThrValPheAlaAlaHisThrAsnLeuAspIleAlaGlnGly 111
OY 572 GTCAACAATGTTGGCTAAAGGCTTGAGAGCTTGACCTCCAGGCGCCATACCTTCC 631
Db 112 ValAsnAsnIleLeuAlaAspLeuHisLeuGlnAsnThrThrMetIleGluLys 131
OY 632 AAAGCTCCCACTACCCCTACAGAGGGAACACCGAGTAGAATTCACGTACATACAC 691
Db 132 TyrSerGluProLysCysLysIleAlaValItyrValProGlu----- 145
OY 692 CAAGAGCTGGACAAATGCTATGTCGAGTG-----AAAGAAATGACGGTGT 739
Db 146 AsnGluLeuGlnSerValArgLeuAlaLeuValAsnGlnValAlaGlnIleGlyThr 165
OY 740 TCTGTCACTCTTTCTCTGCTAGACTGCT-----AATGAGAACAAACAGGATTAAT 793
Db 166 GluTyrThrGluCysThrPheHisThrIleGlyIleGlySerPheLysProGlyAlaAsn 185
OY 794 CTGAATTTACT-----CAGAGGCTTGATGATGATGATGATGATGATGATGATGAT 844
Db 186 AlaAsnProThrIleGluLysAspAlaLeuThrSerValProGluVal-----Lys 203
OY 845 AACCAACAATTTATTCAGAG-----ACGAAATTCGTCA----- 880
Db 204 IleGluAlaIlePheProGlnItyrLeuThrGluThrIleThrLysAlaValLysIleAla 223
OY 881 -----CTGAGAGAGCTTGGCT---CTACATACT----- 907
Db 224 HisProTyrGluGlnProAlaIleAspValItyrThrLeuGlnMetGlnThrItyrLysGlu 243
OY 908 GGAATGGAGAGGTTATGACACATGATGATGATGATGATGATGATGATGATGATGATG 967
Db 244 GlyLeuGlnArgValAlaIleMetLeuProLysLysLeuGlnMetValSerPheIleAsp 263
OY 968 ATAAAAAGACACCTAAACTATCTCATATTCCTTATGCTTGGGCTGGGAGAACCTTA 1027
Db 264 LeuLysThrAlaPheAlaIleAspAsnValArgPhe-----IleGly---AspLeu 279
OY 1028 GAGTCTCAAGTCAAAAGCTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1081
Db 280 LysThrThrValHisLysValAlaIleIleGlyLysAspLysAsnLysPheIleHisGln 299
OY 1082 ---CAGGCTGTGAGGCTTATGCTTACACAGAGGAGATGCTCCATCATGATGATGAT 1138
Db 300 AlaLysSerThrGlyAlaAspValPheIleThrGlyAspValItyrThrHisThrGlnHis 319
OY 1139 GATGCTGCTTCCCAAGAAATATGATATCTCTGTGACACAGC-----AACACTGAA 1192
Db 320 AspLeuLeuAlaIleAsnLeuProThrIleAspAlaGlyHisAsnIleGluLysValMet 339
OY 1193 CGAGGCTTCTTCTACCTTGACATATGATGATGATGATGATGATGATGATGATGATG 1252
Db 340 LysGlyTyrLeuLysAsnLysMetGluGlnAlaLysIleLeuAspTyrGluAlaGlu 359

```

```

OY 1253 AATATCTATACAGACTGACAGGAGCGCTTCTAG 1288
Db 360 PheIleValSerGluValAsnThrAspPropheGln 371

RESULT 13
ID Y705_CAMJE STANDARD: PRT: 241 AA.
AC 09PK2:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein Cj0705.
GN Cj0705.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxId=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies K.M., Feltwell T., Holtroyd S.,
RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL139076; CAB72979.1; -.
DR PIR: D81341; D81341.
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NIF3; 1.
DR TIGRFAMs: TIGR00486; TIGR00486; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 241 AA; 27786 MW; 1574F249905B895A CRC64;

Alignment Scores:
Pred. No.: 5,26e-11 Length: 241
Score: 225.00 Matches: 68
Percent Similarity: 37.01% Conservative: 63
Best Local Similarity: 19.21% Mismatches: 95
Query Match: 8.07% Indels: 128
DB: 1 Gaps: 9

US-09-745-506-74 (1-1553) x Y705_CAMJE (1-241)
OY 245 ATGATTTGAGGCTCTCTTCTTCTGATGATGATGATGATGATGATGATGATGATG 304
Db 1 MetLysLeuSerGluLysIleThrAsnPheLeuAspGlnLeuSerPropheAsnIleGlnGlu 20
OY 305 AGTTGGAGCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 364
Db 21 SerThrAspAsnSerGlyIleLeuLeuGlyAspArgAspSerGlu---IleSerThrVal 39
OY 365 TTCCTGACCAATGCTACTGCTGAGGAGATGATGATGATGATGATGATGATGATGATG 424
Db 40 TyrLeuSerLeuAspIleAspGluValAsnIleIleLysGlnLaseGluAsn-----Ser 57
OY 425 CTCATCTCTCTTACATCCGCTTATCTTCCAGCCATGAGAGGCAATACCTTGAAACACA 484
Db 58 LeuIleIleThrHisHisProLeuIlePheLysGlyLeuLysAspLeuTyrAspLysThr 77

```

QY 485 TGAAGAGAGCGCTGTCATCCGGCTTCGAGAGACAGTCGGTATCTACTCTCCAT 544
 Db 78 TyrProArGalAlaPheIleIleYsGlMetIleYrIysAsnIleSerLeuIleSerMetHis 97
 QY 545 ACACCCATATGCTGCTGCCGCCAGGCGCTCAACAACACTGGTTGGCTTAAGGCTTGAGCT 604
 Db 98 ThrAsnIlyAspLeuSer----- 103
 QY 605 TGTACCTCCAGGCCCATATCATCTTCCAAAGCTCCCACTACCTACAGAGGAAACAC 664
 Db 103 ----- 103
 QY 665 CGAGTAGAATTCAGTTACTACACCCAGACCTGGACAAAGTCATGTCAGCTGAAA 724
 Db 104 -----HisLeuAsnThrTyrPheThrGlu----- 112
 QY 725 GGAATGACGGTGTTCCTGCTACTCTTTCTGCTAGAGACTGGTAATGAGAACAAACA 784
 Db 113 -----IleLeuGlyPheLys----- 117
 QY 785 CGGATTAATCTGAATTTGACTCAGAGGCTTGAATGAGATTCTTCTCCGG 844
 Db 117 ----- 117
 QY 845 AACAAACACTTTCATCAGAGACGGAATTTCTGCTACTGAGAGACCTTGTCTCATAT 904
 Db 118 -----IleSerPheLysAspLysPheLeuIleYr 127
 QY 905 ACTGGAATGGAGCGGTATGCACTGATGATGATGCTGCTCCCTGGACACATGATGAT 964
 Db 128 -----ValGluAsnSerMetSerPheGluAlaLeuLysAsp 139
 QY 965 CGAATAAAGACACCTAAACTATCTCATATTCCTTACGCTTGGGCTGGGGAAC 1024
 Db 140 TrpAlaIlyLysLysIleLeuLeuGlnIleLeuArg-----ValSerAspLys 155
 QY 1025 TTAGAGTTCAGATCAAGTCAGTGTGCGCTGTGCTGCTTCTGGAGACGCTTTCGAG 1084
 Db 156 GlyLysLysAspIleLysArgIleAlaIleCysThrGlySerGlyLysPheLeuIleSer 175
 QY 1085 GGTGTTGAGCTTACCTTACCTACAGATGATGATGATGATGATGATGATGATGATGAT 1144
 Db 176 LysValAspAlaAspCysPheLeuSerGlyAspPheLysTyrHisGlnAlaLeuLysAla 195
 QY 1145 GCTTCCCAAGGAATGAATGTCATCTCTGTGAACACAGCAACAGTGAAGGCTTT-- 1201
 Db 196 LeuSerAsnGlnIleSerLeuIleAspLeuGlyHisPheGluSerGluArgTyrPheSer 215
 QY 1202 -----CTTTCGACCTTCGAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
 Db 216 GlnCysLeuAlaLysAspLeuLysAsnLeu----- 225
 QY 1253 ATTATCTATCAGAGACTGACAGGAGCCCTTCTGAGTGTGA 1294
 Db 226 -----ProLeuGlnValIle 230
 RESULT 14
 YM30_MYCTU STANDARD: PRT; 379 AA.
 ID YM30_MYCTU 010514;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV2230C.
 GN RV2230C OR MT2289 OR MTCY427.11C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NC NCBL_taxid=1773;
 RX SEQUENCE FROM N.A.
 RP STRAIN-H37Rv;

RX MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ernolova M.D., Salzberg S.L.,
 RA Bisher A., Uitterlind T., Weidman J., Khouli H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE UP0135 (NIF3) FAMILY.
 CC -1- CAUTION: REF 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
 CC IN POSITIONS 254 AND 289.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC
 DR EMBL: 270692; CAA94653.1;
 DR EMBL: AE007073; -; NOT_ANNOTATED_CDS.
 DR PIR: B70777; B70777.
 DR TIGR: MT2289; -;
 DR Tuberculist; RV2230C; -;
 DR InterPro; IPR002678; DUF34.
 DR Pfam; PF01784; NIF3; 1.
 DR TIGRFAMS; TIGR00486; TIGR00486; 1.
 KW Hypothetical protein, complete proteome.
 SO SEQUENCE 379 AA; 39397 MW; F4B5E2397426C3F3 CRC64;
 Alignment Scores:
 Pred. No.: 3,22e-10 Length: 379
 Score: 216.50 Matches: 95
 Percent Similarity: 44.01% Conservative: 63
 Best Local Similarity: 26.46% Mismatches: 158
 Query Match: 7.76% Indels: 43
 Gaps: 15
 US-09-745-506-74 (1-1553) x YM30_MYCTU (1-379)
 QY 299 GCTGAGAGTGGGAGAAATGTTGATTCCTGCGGAGAACCAAGCCACACATCTGAAT 358
 Db 21 AlaGlnSerTyrPaspSerValGly--LeuValCysGlyAspProAspAspValAlaAsp 39
 QY 359 ACACCTTCCTGACCAATGACCTGACCTGAGAGAGATGAGAGAGCTGCTGCAAAAGA 418
 Db 40 SerValThrValAlaValAlaAspAlaThrProAlaValAlaAspGlnValProGln----- 57
 QY 419 CGAGACTCATCTCTCTCTACATCCGCTATCTTCGACACCCATGAGAGAGGATGAGAGAG 478
 Db 58 AlaGlyLeuLeuValHisHisProLeuLeuArgGlyValAlaPheThrValAlaAla 77
 QY 479 AACACATGGAAGAGAGCGCTGATCCGGGCTCTGGAACACAGAGTGGATCTACTCT 538
 Db 78 AsnThrProLysGlyValLeuValHisArgLeuIleArgThrGlyArgSerLeuPheThr 97
 QY 539 CCTCATACAGCTATGATGCTCGCCGCCAGAGGCGTCAACACACTGTTGGCTAAAGGCTT 598

```

Db      98  AlAHISTHrAsnAlaSPSerAlaSerProGlyValSerSerAlaLeuAlaAlaVal 117
QY      599  GGAGCTTGTACSTCCAGGCCATACATCTTCCAAAGCTCCCAAC-----643
Db      116  GtyleuthrValasprAlaValLeuAspPro-----ValProGlyAlaAlaAspLeuAsp 135
QY      644  -----TACCCTACAGAGGAAACCACCGAGTA-----GAA 673
Db      136  LysTrpValIleTyrValProArgLysLysSerGluAlaValAlaArgAlaValAlaPheGlu 155
QY      674  TTCAACGTTAACTACACCCCAAGACCTGGACAAGTCAATGCTGCACAGTGAAGAAATT---730
Db      156  AlAGlYAlaGlyAlaIstIleGlyAspTyrSerHisCysSerTrpPheValAlaGlyThrGly 175
QY      731  -----GACGGTGTCTGTCACCTTTCTTCTGCTAGATGGATGATGAG 775
Db      176  GluPheLeuAlaHisAspGlyAlaSerProAlaIleGlySerValGlyThrValGlyArg 195
QY      776  ---GAACAAACAGCGATTATCTGAATTTGACTCAGACGCTTGAATGACGAGTGTAGAT 832
Db      196  ValAlaGluAspArgValGlyValAlaValAlaProAlaArgAlaArgAlaGlyValIleAla 215
QY      833  TTTCTTCCCGGAACAACAACCTTTATCAGAAAGACG-----GAAATTCTGCACCTGGAG 886
Db      216  AlAmet---ArgAlaAlaHisProTyrGluGluProAlaPheAspIlePheAlaLeuVal 234
QY      887  AACGCTTGTCTGCTACACTAGTGAATGGACGCTTATGCACTGCATGATGATGCTCTCC 946
Db      235  ProPro---ProValIleSerGlyLeuGlyArgIleGlyAlaGlyProLysProGluPro 253
QY      947  CTGGCAACCATGATGATGATGAATAAAGACACCTTAACCTATCATATGCTGATGCC 1006
Db      254  LeuArgThrPheValAlaArgLeuGluAlaAlaLeuProThr-----Ala 269
QY      1007  CTTGGGGGGGAGAAACCTTAGAGTCAAGTCAAAATC-----GTGGCCCTGTGTGCT 1060
Db      270  ThGlyValAlaArgAlaAlaGlyAspProAspLeuValSerArgValAlaValCysGly 289
QY      1061  GGTTCGGGAGGACGCTTCTGCAGCGGTGTGAGGCT-----GACCTTACCTCACA 1111
Db      290  GlyAlaGlyAspSerLeuLeuAlaThrValAlaAlaAspValAlaAlaValThr 309
QY      1112  GGTGAGATGTCCATCAT-----GATACTTTGATGCTGCTCCCAAGGAATAAATGTC 1165
Db      310  AlaAspLeuArgHisHisProAlaAspGluHisCysArgAlaSerGln---ValAlaLeu 328
QY      1166  ATCTCTCTGTGAACAACAACACTGAAGAGGCTTTCTTTCTGACCTTGGAGATATGCTG 1225
Db      329  IleAspValAlaHisTrpAlaSerGluPheProTyrCysGlyGlnAlaAlaGlyValLeu 348
QY      1226  GATTCCTCACTTGGAGATAAGATAATATATTCCTATCAGAGACTGACAGGAGCCT 1282
Db      349  ArgSerHisPheGlyAlaSerLeuProValArgValCysThrIleCysThrAspPro 367

RESULT 15
Y931_STRPY STANDARD: PRT: 262 AA.
AC 09A049:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein SPY0931/SPYM3_0644/SPS1208/spym18_0988.
GN SPY0931 OR SPYM3_0644 OR SPS1208 OR SPYM18_0988.
OS Streptococcus pyogenes (serotype M3), and
OS Streptococcus pyogenes (serotype M18).
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466, 186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;

```

```

RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian X., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Tian X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS315 / Serotype M3;
RX MEDLINE=22139808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus;
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SST-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes: a comparative analysis
RT of S. pyogenes SST-1, SF370 and MGAS8232.";
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS8232 / Serotype M18;
RX MEDLINE=1927593; PubMed=11917108;
RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [5]
RP SIMILARITY: BELONGS TO THE UPF0135 (NMF3) FAMILY.
CC - SIMILARITY: BELONGS TO THE UPF0135 (NMF3) FAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC CC
DR EMBL: AE006542; MAK3846.1; -
DR EMBL: AE014149; AAM79251.1; -
DR EMBL: AP005144; BAC64303.1; -
DR EMBL: AE010027; AAL97628.1; -
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; NTF3; 1.
DR TIGRfam: TIGR00486; TIGR00486; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 262 AA; 29289 MW; 583B6720E34D2F29 CRC64;

Alignment Scores:
Pred. No.: 9, 96e-10 Length: 262
Score: 210.00 Matches: 71
Percent Similarity: 38.53% Conservative: 65
Best Local Similarity: 20.11% Mismatches: 115
Query Match: 7.53% Indels: 102
DB: 1 Gaps: 11

US-09-745-506-74 (1-1553) x Y931_STRPY (1-262)
QY 245  ATGGAATTGAAGCTCTCTTCTTCTGGAATGACTTTCATCCCTCTGTTGCTGAG 304
Db      1  MetLysAlaIysThrLeuIleAspAlaTyrGluAlaPheCysProLeuAspLeuSerMet 20

```


QY	305	AGTTGGGCAATGTTGGATTACTGTGGTAACCAAGGCCAACCATACTAGTAAATACATC	366
DB	21	GIuGlYsApValIySGIyLeuGlnMet---GlySerLeuAspIyAspIleAlYVal	39
QY	365	TTTCCTGACCAATGACCTGCATGTAGCAAGATGTAGGAGAGGTCTGCAAAAGMAGGACAC	424
DB	40	MetIleThrLeuAspIleArgGluSerThrValAlaGluAlaIleIyAsnGluValAsp	59
QY	425	CTCATTTCTCTCTACACCCGCTATCTCTCGACCCCATGAAGGCCATTAACCTGGAAACA	484
DB	60	LeuIleIleThrLysHsAlaProIlePheLysProLeuLysAspLeuValSerSerPro	79
QY	485	TGGAAGAGAGCCCTGGTATCCGGGCTCTGGAGAACAAGATCGGTATCTACTCTCCAT	544
DB	80	GlnItrGAsp---IleLeuLeuAspLeuValLysHsAspIleSerValTyValSerHis	98
QY	545	ACAGCCTATGATGCTGCGCCGCCAGGCGCTCAACAACATGCTTGTTGGCTAAAGGCTTGAGCT	604
DB	99	ThrsAsnIleAspIleValProGlyGlyLeuAsnAspTrpPheCys-----	113
QY	605	TGTACCTCCAGGCCCATACATCTTTCCAAAGCTCCAACTCCAACTACAGAGGAAACAC	664
DB	113	-----	113
QY	665	CGAGTAGAATTCAACGTTTAACTAACCCAGACCTGGCAAAAGTCATGTCGACGTGAA	724
DB	114	-----AspLeuGlnIleIySGluAlaThr---	122
QY	725	GGAATTGACGGGTTCCTGTCACCTCTTTCTCTGTAGACTGGTAATGAGGAACAACA	784
DB	122	-----	122
QY	785	CGGATTAACTGAAATTGACTACAGAAAGCTTTGATGACAGTGTAGATTTCTTTCCCG	844
DB	123	-----TyrLeuSerCln	126
QY	845	AACAACAACATTATCAGAAAGCAAAATTCTGTACACTGAGAAAGCTTTCCTCTACAT	904
DB	127	ThrLysGluGlyPhe-----	131
QY	905	ACTGGAAGGGGCGGTATGCACACTGTGATGAATGCTGTCCGTGGCAACCATGATGAT	964
DB	132	---GlyIleGlyAlaGlyIleGlyThyValLysGluGln---AlaLeuGluGluLeuAlaSer	149
QY	965	CGAATTAAGAAGCACACTAAACTATATCTATATCTGCTTAAGCCCTTGGGTGGGGAGAAC	1024
DB	150	LysValLysArgValPheAspLeuAspThrValArgLeu-----IleArgTyAsp	166
QY	1025	TTTAGAGTCAAA---GTCAAAAGTCTGGCCCTGTGTGCTGTCTGGGAGGACGCTTGTG	1081
DB	167	LysGluAsnProLeuIleSerLysIleAlaIleCysGlyGlySerGlyGlyLeuPheTy	186
QY	1082	CAGGCTGTT-----GAGCGTACCTTAACTCACAGAGGAGATATCCCATCATGAT	1132
DB	187	GlnAspAlaValGlnLysGlyAlaAspValTyIleThrGlyAspIleTyThrHisThr	206
QY	1133	ACTTTGATGCTGCTCTCCAAAGAAATAATGTACATCTGTGAAACAGACAACTGAA	1192
DB	207	AlaGlnGluMetLeuThrGluGlyLeuPheAlaValAspProGlyHis---HisIleGlu	225
QY	1193	CGAGGCTTTCCTTTCGACCTTCGAGATATGCTGGATTTCTACCTTGGACAATAG-----	1248
DB	226	ValLeuPheThrGluLysLeuLysGluLysLeuGlnGlyTrpLysGluGluAsnGlyTrp	245
QY	1247	---ATAAATATATATCTATCAGAGCTGACAGGAGCGCT	1282
DB	246	AspValSerIleIleIleSerSerLysLysAspThrAsnPro	258

Search completed: August 22, 2003, 14:00:24
Job time : 45.5 secs